An Interop Publication

CONNEXIONS

The Interoperability Report

April 1995 Volume 9, No. 4

ConneXions — The Interoperability Report tracks current and emerging standards and technologies within the computer and communications industry.

In this issue:

ConneXions is published monthly by Interop Company, a division of SOFTBANK Exposition and Conference Company, 303 Vintage Park Drive, Foster City, California, 94404–1138, USA.
Phone: +1 (415) 578-6900
Fax: +1 (415) 525-0194
E-mail: connexions@interop.com

Subscription hotline: 1-800-575-5717
or +1-502-493-3217

ISSN 0894–5926

From the Editor

The World-Wide Web continues to be the most popular and the fastest growing application on the Internet. While "cruising the Web" might be a fun thing to do in your spare time, it isn't necessarily the most efficient way to find what you are looking for. Enter Web robots, programs that connect to Web sites and collect information for you "while you sleep." In our first article this month, Martijn Koster of NEXOR Ltd. describes how Web robots work and discusses the advantages and disadvantages of robots, with an emphasis on robots used for resource discovery.

History will be written this month when the NSFNET backbone is disconnected and replaced by the interconnected mesh of Internet Service Providers. It doesn't seem that long ago that we published an article entitled "The New NSFNET Backbone Network," (December 1988 actually—*ConneXions* is coming of age too). It seems appropriate at this time to look back and review some of the events that brought us to where we are today. We asked John Quarterman from Texas Internet Consulting to write us an "Internet timeline." The history of the Internet is closely linked to the evolution of the computer networking industry, and you can get some idea of its growth by looking at the Interop (and later NetWorld+Interop) attendance figures scattered throughout the article. In reading John's piece, it also struck me that I could have provided references to many *ConneXions* articles that deal with the events listed in the timeline, but rather than take up more space I will refer you to the 1987–1994 index booklet or the cumulative ASCII index found at the URL: `http://www.interop.com`

The work of the Security Area of the *Internet Engineering Task Force* (IETF) has been described previously in this journal. Jim Galvin of Trusted Information Systems gives us another update from the December 1994 IETF meeting which was held in San Jose, California. The IETF is the primary standards development body for the Internet, and if you are interested in participating in its work you should make note of the information at the end of Dr. Galvin's article.

After a long but unintentional pause, Book Reviews are back. There is also one lone Letter to the Editor. We're always eager to hear from you, so please send us your comments, suggestions or questions. Prospective authors should also contact us in order to receive submission guidelines. The simplest way to reach us is via e-mail to: `connexions@interop.com`

Robots in the Web: threat or treat?

by Martijn Koster, NEXOR Ltd.

Abstract

Robots have been operating in the World-Wide Web for over a year. In that time they have performed useful tasks, but also on occasion wreaked havoc on the networks. This article investigates the advantages and disadvantages of robots, with an emphasis on robots used for resource discovery. New alternative resource discovery strategies are discussed and compared. It concludes that while current robots will be useful in the immediate future, they will become less effective and more problematic as the Web grows.

Introduction

The World-Wide Web [1] has become highly popular in the last few years, and is now one of the primary means of information publishing on the Internet. When the size of the Web increased beyond a few sites and a small number of documents, it became clear that manual browsing through a significant portion of the hypertext structure is no longer possible, let alone an effective method for resource discovery.

This problem has prompted experiments with automated browsing by "robots." A Web robot is a program that traverses the Web's hypertext structure by retrieving a document, and recursively retrieving all documents that are referenced. These programs are sometimes called "Spiders," "Web Wanderers," or "Web Worms." These names, while perhaps more appealing, may be misleading, as the term "spider" and "wanderer" give the false impression that the robot itself moves, and the term "worm" might imply that the robot multiplies itself, like the infamous Internet worm [2]. In reality robots are implemented as a single software system that retrieves information from remote sites using standard Web protocols.

Robot uses

Robots can be used to perform a number of useful tasks:

- *Statistical Analysis:* The first robot [3] was deployed to discover and count the number of Web servers. Other statistics could include the average number of documents per server, the proportion of certain file types, the average size of a Web page, the degree of interconnectedness, etc.

- *Maintenance:* One of the main difficulties in maintaining a hypertext structure is that references to other pages may become "dead links," when the page referred to is moved or even removed. There is currently no general mechanism to proactively notify the maintainers of the referring pages of this change. Some servers, for example the CERN HTTPD, will log failed requests caused by dead links, along with the reference of the page where the dead link occurred, allowing for post-hoc manual resolution. This is not very practical, and in reality authors only find that their documents contain bad links when they notice themselves, or in the rare case that a user notifies them by e-mail.

  A robot that verifies references, such as MOMspider [4], can assist an author in locating these dead links, and as such can assist in the maintenance of the hypertext structure. Robots can help maintain the content as well as the structure, by checking for HTML [5] compliance, conformance to style guidelines, regular updates, etc., but this is not common practice. Arguably this kind of functionality should be an integrated part of HTML authoring environments, as these checks can then be repeated when the document is modified, and any problems can be resolved immediately.

- *Mirroring:* Mirroring is a popular technique for maintaining FTP archives. A mirror copies an entire directory tree recursively by FTP, and then regularly retrieves those documents that have changed. This allows load sharing, redundancy to cope with host failures, and faster and cheaper local access, and off-line access.

  In the Web, mirroring can be implemented with a robot, but at the time of writing no sophisticated mirroring tools exist. There are some robots that will retrieve a subtree of Web pages and store it locally, but they don't have facilities for updating only those pages that have changed. A second problem unique to the Web is that the references in the copied pages need to be rewritten: where they reference pages that have also been mirrored they may need to changed to point to the copies, and where relative links point to pages that haven't been mirrored they need to be expanded into absolute links. The need for mirroring tools for performance reasons is much reduced by the arrival of sophisticated caching servers [6], which do offer selective updates, can guarantee that a cached document is up-to-date, and are largely self maintaining. However, it is expected that mirroring tools will be developed in due course.

- *Resource discovery:* Perhaps the most exciting application of robots is their use in resource discovery. Where humans cannot cope with the amount of information it is attractive to let the computer do the work. There are several robots that summarise large parts of the Web, and provide access to a database with these results through a search engine.

  This means that rather than relying solely on browsing, a Web user can combine browsing and searching to locate information; even if the database doesn't contain the exact item you want to retrieve, it is likely to contain references to related pages, which in turn may reference the target item.

  The second advantage is that these databases can be updated automatically at regular intervals, so that dead links in the database will be detected and removed. This in contrast to manual document maintenance, where verification is often sporadic and not comprehensive. The use of robots for resource discovery will be further discussed below.

- *Combined Uses:* A single robot can perform more than one of the above tasks. For example the RBSE Spider [7] does statistical analysis of the retrieved documents as well providing a resource discovery database. Such combined uses are unfortunately quite rare.

Operational costs and dangers

The use of robots comes at a price, especially when they are operated remotely on the Internet. In this section we will see that robots can be dangerous in that they place high demands on the Web.

- *Network resource and server load:* Robots require considerable bandwidth. Firstly robots operate continually over prolonged periods of time, often months. To speed up operations many robots feature parallel retrieval, resulting in a consistently high use of bandwidth in the immediate proximity. Even remote parts of the network can feel the network resource strain if the robot makes a large number of retrievals in a short time ("rapid fire"). This can result in a temporary shortage of bandwidth for other uses, especially on low-bandwidth links, as the Internet has no facility for protocol-dependent load balancing.

continued on next page

Robots in the Web *(continued)*

Traditionally the Internet has been perceived to be "free," as the individual users did not have to pay for its operation. This perception is coming under scrutiny, as especially corporate users do feel a direct cost associated with network usage. A company may feel that the service to its (potential) customers is worth this cost, but that automated transfers by robots are not.

Besides placing demands on network, a robot also places extra demand on servers. Depending on the frequency with which it requests documents from the server this can result in a considerable load, which results in a lower level of service for other Web users accessing the server. Especially when the host is also used for other purposes this may not be acceptable. As an experiment, the author ran a simulation of 20 concurrent retrievals from his server running the Plexus software on a Sun 4/330. Within minutes the machine slowed down to a crawl and was not usable for anything. Even with only consecutive retrievals the effect can be felt. Only the week that this article was written a robot visited the author's site with rapid fire requests. After 170 consecutive retrievals the server, which had been operating fine for weeks, crashed under the extra load.

This shows that rapid fire needs to be avoided. Unfortunately even modern manual browsers (e.g., Netscape) contribute to this problem by retrieving in-line images concurrently. The Web's protocol, HTTP [8], has been shown to be inefficient for this kind of transfer [9], and new protocols are being designed to remedy this [10].

- *Updating overhead:* It has been mentioned that databases generated by robots can be automatically updated. Unfortunately there is no efficient change control mechanism in the Web; There is no single request that can determine which of a set of URLs has been removed, moved, or modified.

  HTTP does provide the "If-Modified-Since" mechanism, whereby the user-agent can specify the modification time-stamp of a cached document along with a request for the document. The server will then only transfer the contents if the document has been modified since it was cached.

  This facility can only be used by a robot if it retains the relationship between the summary data it extracts from a document, its URL, and the timestamp of the retrieval. This places extra requirements on the size and complexity on the database, and is not widely implemented.

- *Client-side robots/agents:* The load on the network is especially an issue with the category of robots that are used by end-users, and implemented as part of a general purpose Web client (e.g., the Fish Search [11] and the TkWWW robot [12]).

  One feature that is common in these end-user robots is the ability to pass on search-terms to search engines found while traversing the Web. This is touted as improving resource discovery by querying several remote resource discovery databases automatically. However it is the author's opinion that this feature is unacceptable for two reasons. Firstly, a search operation places a far higher load on a server than a simple document retrieval, so a single user can cause a considerable overhead on several servers in a far shorter period than normal.

Secondly, it is a fallacy to assume that the same search-terms are relevant, syntactically correct, let alone optimal for a broad range of databases, and the range of databases is totally hidden from the user. For example, the query "Ford and garage" could be sent to a database on 17th century literature, a database that doesn't support Boolean operators, or a database that specifies that queries specific to automobiles should start with the word "car:". And the user isn't even aware of this.

Another dangerous aspect of a client-side robot is that once it is distributed no bugs can be fixed, no knowledge of problem areas can be added and no new efficient facilities can be taken advantage of, as not everyone will upgrade to the latest version.

The most dangerous aspect however is the sheer number of possible users. While some people are likely to use such a facility sensibly, i.e., bounded by some maximum, on a known local area of the Web, and for a short period of time, there will be people who will abuse this power, through ignorance or arrogance. It is the author's opinion that remote robots should not be distributed to end-users, and fortunately it has so far been possible to convince at least some robot authors to cancel releases [13].

Even without the dangers, client-side robots pose an ethical question: where the use of a robot may be acceptable to the community if its data is made available to the community, client-side robots may not be acceptable as they operate only for the benefit a single user. The ethical issues will be discussed further below.

End-user "intelligent agents" [14] and "digital assistants" are currently a popular research topic in computing, and often viewed as the future of networking. While this may indeed be the case, and it is already apparent that automation is invaluable for resource discovery, a lot more research is required for them to be effective. Simplistic user-driven Web robots are far removed from intelligent network agents: an agent needs to have some knowledge of where to find specific kinds of information (i.e., which services to use) rather than blindly traversing all information. Compare the situation where a person is searching for a bookshop; they use the Yellow Pages for a local area, find the list of shops, select one or a few, and visit those. A client-side robot would walk into *all* shops in the area asking for books. On a network, as in real life, this is inefficient on a small scale, and prohibitive on a larger scale.

- *Bad Implementations:* The strain placed on the network and hosts is sometimes increased by bad implementations of especially newly written robots. Even if the protocol and URLs sent by the robot is correct, and the robot correctly deals with returned protocol (including more advanced features such as redirection), there are some less-obvious problems.

  The author has observed several identical robot runs accessing his server. While in some cases this was caused by people using the site for testing (instead of a local server), in some cases it became apparent that this was caused by lax implementation. Repeated retrievals can occur when either no history of accessed locations is stored (which is unforgivable), or when a robot does not recognise cases where several URL are syntactically equivalent, e.g., where different DNS aliases for the same IP address are used, or where URLs aren't canonicalised by the robot, e.g., `foo/bar/../baz.html` is equivalent to `foo/baz.html`.

continued on next page

Robots in the Web *(continued)*

Some robots sometimes retrieve document types, such as GIFs and *PostScript,* which they cannot handle and thus ignore.

Another danger is that some areas of the Web are near-infinite. For example, consider a script that returns a page with a link to one level further down. This will start with for example `/cgi-bin/pit/`, and continue with `/cgi-bin/pit/a/`, `/cgi-bin/pit/a/a/`, etc. Because such URL spaces can trap robots that fall into them, they are often called "black holes." See also the discussion of the Proposed Standard for Robot Exclusion below.

Cataloguing issues

That resource discovery databases generated by robots are popular is undisputed. The author himself regularly uses such databases when locating resources. However, there are some issues that limit the applicability of robots to Web-wide resource discovery.

- *There is too much material, and it's too dynamic:* One measure of effectiveness of an information retrieval approach is "recall," the fraction of all relevant documents that were actually found. Brian Pinkerton [15] states that recall in Internet indexing systems is adequate, as finding enough relevant documents is not the problem. However, if one considers the complete set of information available on the Internet as a basis, rather than the database created by the robot, recall cannot be high, as the amount of information is enormous, and changes are very frequent. So in practice a robot database may not contain a particular resource that is available, and this will get worse as the Web grows.
- *Determining what to include/exclude:* A robot cannot automatically determine if a given Web page should be included in its index. Web servers may serve documents that are only relevant to a local context (for example an index of an internal library), that exists only temporarily, etc. To a certain extent the decision of what is relevant also depends on the audience, which may not have been identified at the time the robot operates. In practice robots end up storing almost everything they come come across. Note that even if a robot could decide if a particular page is to be exclude form its database they have already incurred the cost of retrieving the file; a robot that decides to ignore a high percentage of documents is very wasteful.

  In an attempt to alleviate this situation somewhat, the robot community has adopted "A Standard for Robot exclusion" [16]. This standard describes the use of a simple structured text file available at well-known place on a server (`/robots.txt`) to specify which parts of their URL space should be avoided by robots (see Figure 1). This facility can also be used to warn robots for black holes. Individual robots can be given specific instructions, as some may behave more sensibly than others, or are known to specialise in a particular area. This standard is voluntary, but is very simple to implement, and there is considerable public pressure for robots to comply.

  Determining how to traverse the Web is a related problem. Given that most Web servers are organised hierarchically, a breadth-first traversal from the top to a limited depth is likely to more quickly find a broader and higher-level set of document and services than a depth-first traversal, and is therefore much preferable for resource discovery.

However, a depth-first traversal is more likely to find individual users' home pages with links to other, potentially new, servers, and is therefore more likely to find new sites to traverse.

```
# /robots.txt for http://www.site.com/

User-agent: *                  # attention all robots:
Disallow:   /cyberworld/map    # infinite URL space
Disallow:   /tmp/              # temporary files
```

Figure 1: An example `robots.txt` file

- *Summarising documents:* It is very difficult to index an arbitrary Web document. Early robots simply stored document titles and anchor texts, but newer robots use more advanced mechanisms and generally consider the entire content.

  These methods are good general measures, and can be automatically applied to all Web pages, but cannot be as effective as manual indexing by the author. HTML provides a facility to attach general meta information to documents, by specifying a META element e.g., `<meta name= "Keywords" value= "Ford Car Maintenance">`. However, no semantics have (yet) been defined for specific values of the attributes of this tag, and this severely limits its acceptance, and therefore its usefulness.

  This results in a low "precision," the proportion of the total number of documents retrieved that is relevant to the query. Advanced features such as Boolean operators, weighted matches like WAIS, or relevance feedback can improve this, but given that the information on the Internet is enormously diverse, this will continue to be a problem.

- *Classifying documents:* Web users often ask for a "subject hierarchy" of documents in the Web. Projects such as GENVL [17] allow these subject hierarchies to be manually maintained, which presents a number of problems that fall outside the scope of this article. It would be useful if a robot could present a subject hierarchy view of its data, but this requires some automated classification of documents [18].

  The META tag discussed above could provide a mechanism for authors to classify their own documents. The question then arises which classification system to use, and how to apply it. Even traditional libraries don't use a single universal system, but adopt one of a few, and adopt their own conventions for applying them. This gives little hope for an immediate universal solution for the Web.

- *Determining document structures:* Perhaps the most difficult issue is that the Web doesn't consist of a flat set of files of equal importance. Often services on the Web consist of a collection of Web pages: there is a welcome page, maybe some pages with forms, maybe some pages with background information, and some pages with individual data points. The service provider announces the service by referring to the welcome page, which is designed to give structured access to the rest of the information.

continued on next page

Robots in the Web *(continued)*

A robot however has no way of distinguishing these pages, and may well find a link into for example one of the data points or background files, and index those rather than the main page. So it can happen that rather than storing a reference to "The Perl FAQ," it stores some random subset of the questions addressed in the FAQ. If there was a facility in the Web for specifying—per document—that someone shouldn't link to the page, but to another one specified, this problem could be avoided.

Related to the above problem is that the content of Web pages are often written for a specific context, provided by the access structure, and may not make sense outside that context. For example, a page describing the goals of a project may refer to "The project," without fully specifying the name, or giving a link to the welcome page. Another problem is that of moved URLs. Often when service administrators reorganise their URL structure they will provide mechanisms for backward compatibility with the previous URL structure, to prevent broken links. In some servers this can be achieved by specifying redirection configuration, which results in the HTTP negotiating a new URL when users try to access the old URL. However, when symbolic links are used it is not possible to tell the difference between the two. An indexing robot can in these cases store the deprecated URL, prolonging the requirement for a Web administrator to provide backward compatibility.

A related problem is that a robot might index a mirror of a particular service, rather than the original site. If both source and mirror are visited there will be duplicate entries in the database, and bandwidth is being wasted repeating identical retrievals to different hosts. If only the mirror is visited users may be referred to out-of-date information even when up-to-date information is available elsewhere.

Ethics

We have seen that robots are useful, but that they can place high demands on bandwidth, and that they have some fundamental problems when indexing the Web. Therefore a robot author needs to balance these issues when designing and deploying a robot. This becomes an ethical question "Is the cost to others of the operation of a robot justified?" This is a grey area, and people have very different opinions on what is acceptable.

When some of the acceptability issues first became apparent (after a few incidents with robots doubling the load on servers) the author developed a set of Guidelines for Robot Writers [19], as a first step to identify problem areas and promote awareness. These guidelines can be summarised as follows:

- Reconsider: Do you really need a new robot?
- Be accountable: Ensure the robot can be identified by server maintainers, and the author can be easily contacted.
- Test extensively on local data
- Moderate resource consumption: Prevent rapid fire and eliminate redundant and pointless retrievals.
- Follow the Robot Exclusion Standard.
- Monitor operation: Continuously analyse the robot logs.
- Share results: Make the robot's results available to others, the raw results as well as any intended high-level results.

David Eichman [20] makes a further distinction between *Service Agents,* robots that build information bases that will be publicly available, and *User Agents,* robots that benefit only a single user such as client-side robots, and has identified separate high-level ethics for each.

The fact that most robot writers have already implemented these guidelines indicates that they are conscious of the issues, and eager to minimise any negative impact. The public discussion forum provided by the robots mailing list speeds up the discussion of new problem areas, and the public overview of the robots on the Active list provides a certain community pressure on robot behaviour [21].

This maturation of the robot field means there have recently been fewer incidents where robots have upset information providers. Especially the standard for robot exclusion means that people who don't approve of robots can prevent being visited. Experiences from several projects that have deployed robots have been published, especially at the World-Wide Web conferences at CERN in July 1994 and Chicago in October 1994, and these help to educate, and discourage, would-be robot writers. However, with the increasing popularity of the Internet in general, and the Web in particular it is inevitable that more robots will appear, and it is likely that some will not behave appropriately.

Alternatives for resource discovery

Robots can be expected to continue to be used for network information retrieval on the Internet. However, we have seen that there are practical, fundamental and ethical problems with deploying robots, and it is worth considering research into alternatives, such as ALIWEB [22] and Harvest [23].

ALIWEB has a simple model for human distributed indexing of services in the Web, loosely based on Archie [24]. In this model aggregate indexing information is available from hosts on the Web. This information indexes only local resources, not resources available from third parties. In ALIWEB this is implemented with IAFA templates [25], which give typed resource information in a simple text-based format (See Figure 2). These templates can be produced manually, or can be constructed by automated means, for example from titles and META elements in a document tree. The ALIWEB gathering engine retrieves these index files through normal Web access protocols, and combines them into a searchable database. Note that it is not a robot, as it doesn't recursively retrieve documents found in the index.

```
Template-Type: SERVICE
Title:         The ArchiePlex Archie Gateway
URL:           /public/archie/archieplex/archieplex.html
Description:   A Full Hypertext interface to Archie.
Keywords:      Archie, Anonymous FTP.
Template-Type: DOCUMENT
Title:         The Perl Page
URL:           /public/perl/perl.html
Description:   Information on the Perl Programming
               Language. Includes hypertext versions
               of the Perl 5 Manual and the latest FAQ.
Keywords:      perl, programming language, perl-faq
```

Figure 2: An IAFA index file

continued on next page

Robots in the Web *(continued)*

There are several advantages to this approach. The quality of human-generated index information is combined with the efficiency of automated update mechanisms. The integrity of the information is higher than with traditional "hotlists," as only local index information is maintained. Because the information is typed in a computer-readable format, search interfaces can offer extra facilities to constrain queries. There is very little network overhead, as the index information is retrieved in a single request. The simplicity of the model and the index file means any information provider can immediately participate.

There are some disadvantages. The manual maintenance of indexing information can appear to give a large burden on the information provider, but in practice indexing information for major services don't change often. There have been experiments with index generation from TITLE and META tags in HTML, but this requires the local use of a robot, and has the danger that the quality of the index information suffers. A second limitation is that in the current implementation, information providers have to register their index files at a central registry, which limits scalability. Finally, updates are not optimally efficient, as an entire index file needs to retrieved even if only one of its records was modified.

ALIWEB has been in operation since October 1993, and the results have been encouraging. The main operational difficulties appeared to be lack of understanding; initially people often attempted to register their own HTML files instead of IAFA index files. The other problem is that as a personal project ALIWEB is run on a spare-time basis and receives no funding, so further development is slow.

Harvest is a distributed resource discovery system recently released by the Internet Research Task Force Research Group on Resource Discovery (IRTF-RD), and offers software systems for automated indexing contents of documents, efficient replication and caching of such index information on remote hosts, and finally searching of this data through an interface in the Web. Initial reactions to this system have been very positive.

One disadvantage of Harvest is that it is a large and complex system which requires considerable human and computing resource, making it less accessible to information providers.

The use of Harvest to form a common platform for the interworking of existing databases is perhaps its most exciting aspect. It is reasonably straightforward for other systems to interwork with Harvest; experiments have shown that ALIWEB for example can operate as a Harvest broker. This gives ALIWEB the caching and searching facilities Harvest offers, and offers Harvest a low-cost entry mechanism.

These two systems show attractive alternatives to the use of robots for resource discovery: ALIWEB provides a simple and high-level index, Harvest provides comprehensive indexing system that uses low-level information. However, neither system is targeted at indexing of third-parties that don't actively participate, and it is therefore expected that robots will continue to be used for that purpose, but in co-operation with other systems such as ALIWEB and Harvest.

Conclusions

In today's World-Wide Web, robots are used for a number of different purposes, including global resource discovery. There are several practical, fundamental, and ethical problems involved in the use of robots for this task.

The practical and ethical problems are being addressed as experience with robots increases, but are likely to continue to cause occasional problems. The fundamental problems limit the amount of growth there is for robots. Alternatives strategies such as ALIWEB and Harvest are more efficient, and give authors and sites control of the indexing of their own information. It is expected that this type of system will increase in popularity, and will operate alongside robots and interwork with them. In the longer term, complete Web-wide traversal by robots will become prohibitively slow, expensive, and ineffective for resource discovery.

References

[1] Berners-Lee, T., R. Cailliau, A. Loutonen, H. F. Nielsen and A. Secret, "The World-Wide Web," *Communications of the ACM,* Volume 37, No. 8, August 1994, pp. 76–82.

[2] Seeley, Donn. "A tour of the worm," USENIX Association Winter Conference 1989 Proceedings, January 1989, pp. 287–304.

[3] Gray, M., "Growth of the World-Wide Web," Dec. 1993. URL: `http://www.mit.edu:8001/aft/sipb/user/mkgray/ht/web-growth.html`

[4] Fielding, R., "Maintaining Distributed Hypertext Infostructures: Welcome to MOMspider's Web," Proc. First International World-Wide Web Conference, Geneva, Switzerland, May 1994.

[5] Berners-Lee, T., D. Conolly at al., "HyperText Markup Language Specification 2.0," Work in progress of the HTML working group of the IETF, Internet Draft.

[6] Luotonen, A., K. Altis, "World-Wide Web Proxies," Proceedings of the First International World-Wide Web Conference, May 1994.

[7] Eichmann, D., "The RBSE Spider—Balancing Effective Search against Web Load," Proceedings of the First International World-Wide Web Conference, May 1994.

[8] Berners-Lee, T., R. Fielding, F. Nielsen, "HyperText Transfer Protocol," Work in progress of the HTTP working group of the IETF, Internet Draft.

[9] Spero, S., "Analysis of HTTP Performance problems," July 1994. URL: `http://sunsite.unc.edu/mdma-release/http-prob.html`

[10] Spero, S., "Progress on HTTP-NG". URL: `http://info.cern.ch/hypertext/www/Protocols/HTTP-NG/http-ng-status.html`

[11] De Bra, P. M. E. and R. D. J. Post, "Information Retrieval in the World-Wide Web: Making Client-based searching feasible," Proc. First International World-Wide Web Conference, May 1994.

[12] Spetka, Scott, "The TkWWW Robot: Beyond Browsing," Proceedings of the Second International World-Wide Web Conference, Chicago, United States, October 1994.

[13] Slade, R., "Risks of client search tools," RISKS-FORUM Digest, Volume 16, No. 37, Weds 31 August 1994.

[14] Riechen, Doug, "Intelligent Agents," *Communications of the ACM,* Volume 37, No. 7, July 1994.

[15] Pinkerton, B., "Finding What People Want: Experiences with the WebCrawler," Proceedings of the Second International World-Wide Web Conference, October 1994.

continued on next page

Robots in the Web *(continued)*

[16] Koster, M., "A Standard for Robot Exclusion," URL: `http://web. nexor.co.uk/mak/doc/robots/norobots.html`

[17] McBryan, A., "GENVL and WWWW: Tools for Taming the Web," Proceedings of the First International World-Wide Web Conference, May 1994.

[18] Kent, R. E., Neus, C., "Creating a Web Analysis and Visualization Environment," Proceedings of the Second International World-Wide Web Conference, October 1994.

[19] Koster, Martijn. "Guidelines for Robot Writers," 1993. URL: `http://web.nexor.co.uk/mak/doc/robots/guidelines.html`

[20] Eichmann, D., "Ethical Web Agents," Proceedings of the Second International World-Wide Web Conference, October 1994.

[21] Koster, Martijn, "WWW Robots, Wanderers and Spiders," URL: `http://web.nexor.co.uk/mak/doc/robots/robots.html`

[22] Koster, Martijn, "ALIWEB Archie-Like Indexing in the Web," Proceedings of the First International World-Wide Web Conference, May 1994.

[23] Bowman, Mic, Peter B. Danzig, Darren R. Hardy, Udi Manber and Michael F. Schwartz, "Harvest: Scalable, Customizable Discovery and Access System," Technical Report CU-CS-732-94, Department of Computer Science, University of Colorado, Boulder, July 1994. URL: `http://harvest.cs.colorado.edu/`

[24] Deutsch, P., A. Emtage, "Archie—An Electronic Directory Service for the Internet," Proceedings of the USENIX Winter Conference, pp. 93–110, January 1992.

[25] Deutsch, P., A. Emtage, M. Koster, and M. Stumpf, "Publishing Information on the Internet with Anonymous FTP," Work in progress of the Integrated Internet Information Retrieval working group of the IETF, Internet Draft.

[26] Mark Handley & Jon Crowcroft, "The World-Wide Web: How Servers Work," *ConneXions,* Volume 9, No. 2, February 1995.

[27] Deutsch, P., "The Internet as Service Provider," *ConneXions,* Volume 7, No. 9, September 1993.

[28] Mark Handley and Jon Crowcroft, *The World Wide Web: Beneath the Surf,* ISBN 1-85728-435-6, UCL Press, 1995.

[29] Berners-Lee, T., "A Summary of the WorldWideWeb System," *ConneXions,* Volume 6, No. 7, July 1992.

[30] Kahle, B., "An Information System for Corporate Users: Wide Area Information Servers (WAIS)," *ConneXions,* Volume 5, No. 11, November 1991.

[31] McCahill, M., "The Internet Gopher: A Distributed Server Information System," *ConneXions,* Volume 6, No. 7, July 1992.

MARTIJN KOSTER holds a B.Sc. in Computer Science from Nottingham University (UK). During his national service he worked on as 2nd lieutenant of the Dutch Army at the Operations Research group of SHAPE Technical Center, NATO's research lab in the Netherlands. Since 1992 he has worked for NEXOR as software engineer on X.500 Directory User Agents, and maintains NEXOR's World-Wide Web service (`http://web.nexor.co.uk/`). He is also author of the ALIWEB, CUSI and ArchiePlex search tools, and maintains a mailing-list dedicated to World-Wide Web robots. He can be reached as: `m.koster@nexor.co.uk`

The History of the Internet and the Matrix

by John S. Quarterman, Texas Internet Consulting

Introduction Today billboards on highways carry URLs for addresses on the Information Highway, 20,000 year old underground French cave paintings appear in color online within a month of their discovery, live visuals from the Space Shuttle are available on the World-Wide Web in real time, and no television network news show is complete without an electronic mail address. The Internet is a household word, and most of the other parts of the other networks in the worldwide Matrix of interconnected networks are collapsing into its gravitational pull. What are the origins of this phenomenon?

Space doesn't permit us to name all the networks, companies, and people, or even countries, who participated in the development of computer networking as we know it today. Instead we present this timeline of selected facts and interpretations, concentrating almost exclusively on North America and Europe. We also avoid the history of conferencing systems, from EIS to AOL, and choose instead to concentrate on distributed networks.

1945 Vannevar Bush publishes an article in *The Atlantic* magazine "As We May Think," about communicating using computers. Nobody actually does this for many years afterwards, since in the time of U.S. President Harry Truman computers were not up to the task.

1957 The Soviet Union under Kruschev launches Sputnik, which is the first artificial satellite to orbit the earth. In the United States, the Eisenhower administration forms the *Advanced Research Projects Agency* (ARPA) to avoid being left behind again.

1961 The Berlin Wall is built, while Kruschev is in the Kremlin and Kennedy is in the White House.

1964 Paul Baran of the RAND Corporation publishes a report proposing a network technology sufficiently decentralized that a network could survive arbitrary loss of links or nodes, as for instance in a nuclear war. Traditional telephone technology involves setting up a dedicated electrical circuit from one end to the other for the life of a data connection. Such circuit switching wouldn't work well when wires were being lost. Instead, a new technology called *packet switching* was invented, to divide data into small pieces (packets) and switch them through different routes to reach their destinations. Such packets are sometimes called *datagrams,* in analogy to telegrams, since each is short and contains a destination address. If one link (wire) or node (switching computer) blew up, data would simply be routed around it.

Sometime in the late 1960s Rudimentary electronic mail on single timesharing computers is invented. Nobody knows who did it first; probably it was reinvented almost every time a new timesharing operating system was developed.

1968 The first packet switching network is implemented at the National Physical Laboratories (NPL) in the United Kingdom. However, this NPL network did not extend beyond the local site.

ARPA funds the *ARPA Computer Network* (ARPANET), partly to permit its researchers to use supercomputers at other sites, thus avoiding buying a supercomputer for every research site. This was called "resource sharing." The ARPANET was also an experiment in packet switching technology.

continued on next page

History of the Internet *(continued)*

The early researchers at UCLA and elsewhere did not necessarily have Baran's cold war scenarios immediately in mind, but those scenarios probably didn't hurt getting the military funding for the project approved.

1969 The first ARPANET node is installed at UCLA, and four nodes communicate from different geographical locations by the end of the year. The ARPANET was thus the first distributed packet switching network. The prime ARPANET contractor at the outset was Bolt Beranek and Newman (BBN) of Cambridge, MA.

ARPA money was channeled through the Navy, and the ARPANET node in Berkeley is considered some sort of submarine warfare project by some activists during the Johnson administration and the Vietnam War. Serious ARPANET research at this time and for a decade afterwards often required a secret clearance. ARPANET has 12 hosts in December 1969.

1970 Ray Tomlinson of BBN, in collaboration with other ARPANET researchers, invents a program to send an electronic mail message across a distributed network (the ARPANET) and sends the first message. Electronic mail was not in the original plans for the ARPANET, nor were other services for communication among people. This new service follows a pattern that will be repeated many times: network users see a need; someone who wants it badly enough and has the technical skills prototypes it; others help develop it; it spreads to others who find it useful; eventually it becomes standardized and everybody uses it.

The first servers for early versions of the *File Transfer Protocol* (FTP) are written. Mail is first widely deployed as a bag on the side of FTP.

Norman Abrahamson and his team at the University of Hawaii implement a network called *Aloha* that foreshadows some of the basic techniques later used in Ethernet, involving as it does broadcasting on a single channel, with collision detection.

1971 ARPANET has 13 hosts in January 1971.

1972 *TELNET* (remote login) is specified. TELNET, FTP, and mail are the main user application protocols of the early ARPANET. The same Big Three applications reign as the most popular for the next twenty years, although others are developed right along.

ARPANET demonstrated at the *International Conference on Computer Communications* (ICCC) and becomes accepted outside its original backers.

Mid 1970s Mail programs invented that could read and delete messages, and compose headers for new messages and replies; sure beat using a text editor to hand-craft all those headers...

1973 ARPANET has 35 hosts in January 1973.

Bob Metcalfe invents Ethernet, an inexpensive and flexible local area network protocol, which is developed by XEROX, later in conjunction with Digital Equipment Corporation. It and other networking technologies ensured that Local Area Networks (LANs) would spread, and some kind of technology would be needed to link these differing approaches to networking together so that people could use them all.

1974 BBN deploys TELENET, a public data network (PDN) based on ARPANET technology.

First distributed file access: the MIT *Incompatible Timesharing System* (ITS) had network transparent file I/O by January 1974.

About this year, mailing lists are invented on the ARPANET.

1975 The ARPANET is so successful that ARPA doesn't consider it research anymore and hands operational authority over to the *Defense Communications Agency* (DCA). (DCA is now know as DISA, the Defense Information Systems Agency.)

ARPANET has 63 hosts in January 1976.

1976 CCITT (*International Consultative Committee on Telephony and Telegraphy*) approves the first guidelines for X.25, a network protocol using virtual circuits (like traditional telephone circuit switching, except paths are buffers and switches, rather than wires and relays). Networks such as SERCnet and EPSS in the United Kingdom start to use X.25.

XEROX establishes the XEROX Internet, using *XEROX Network Services* (XNS), a sophisticated internet protocol suite. XNS was also quite popular outside XEROX for some time, but XEROX never published all its specifications nor encouraged outside input into its specifications, and TCP/IP took over. XNS lives on, however, in Novell *NetWare,* whose protocols are extensions of XNS.

Many other proprietary network protocol suites, such as SNA (IBM's Systems Network Architecture) and DNA (Digital's Digital Network Architecture, better known as DECnet) flourished. The deployed number of nodes using such protocols is probably still huge. But eventually TCP/IP outgrew them.

Some mailing lists, such as `SF-LOVERS` (about science fiction), take on non-work subjects and grow huge despite attempts of some system administrators to stamp out such frivolity. Mailing lists eventually also become the most essential tool of the working groups that specify the Internet protocols.

1977 The Experimental Internet begins, to develop early experimental versions of the TCP/IP protocols, involving researchers such as Vinton G. Cerf, Robert Kahn, David Clark, and Louis Pouzin. These protocols were named for the two most prominent among them. IP is the *Internet Protocol,* which permits communications among different types of underlying networks, such as Ethernets and Token Rings. TCP is the *Transmission Control Protocol,* which provides reliable communications to end users (people or programs) across IP datagrams.

Early versions of OSI (*Open Systems Interconnection*) are promulgated by CCITT, and written by Hubert Zimmermann and others. OSI emphasizes virtual circuits, and X.25 will become the cornerstone protocol of ISO-OSI.

Technical development of both ISO-OSI and TCP/IP was international from the outset.

1978 *UUCP* (UNIX to UNIX CoPy Protocol) invented at AT&T Bell Laboratories and distributed with successive versions of UNIX.

1979 *USENET* (User's Network) was invented by Tom Truscott and Steve Bellovin and deployed between Duke University and the University of North Carolina. USENET *news* was invented in imitation of ARPANET mailing lists.

continued on next page

History of the Internet *(continued)*

Necessity was the mother of invention both in causing schools that did not have ARPANET connections to be the source of this new service, and in refining the mailing lists idea to make it work over the available links, which then used 300 bps dialup modems. Instead of sending a copy of each message to every person that wanted to read it, which would have required sending multiple copies to each participating machine, USENET sends one copy to each machine and only that one copy is stored on that machine, rather than multiple copies in mailboxes for each user.

USENET originally exclusively used UUCP to carry its news messages, but later the USENET news service became widespread over other networks, including BITNET and even in a modified form over FidoNet, plus especially over the Internet.

ACSnet deployed, using the Sydney UNIX Network protocols developed by Piers Dick-Lauder and others at the University of Sydney, Australia.

1980 The ISO *OSI Reference Model* is published. The split between OSI and TCP/IP networkers (roughly proponents of virtual circuits and datagrams) deepens.

1981 *BITNET* (Because It's Time Network) starts to connect computer centers worldwide, using mostly 9600 bps leased lines, which are at this time a cost-effective alternative to ARPANET's 56Kbps links. It uses a protocol called Network Job Entry (NJE) that emulates Holerith punch cards, but it delivers mail and some other services.

CSNET (Computer Science Network) established to provide ARPANET-like services to computer science departments without ARPANET access to due lack of related government grants. CSNET mostly provides mail access through a specialized dialup protocol, but also provides IP connectivity over X.25 links to some sites.

Unlike USENET, which was invented and deployed by graduate students with no official authorization or funding, CSNET is designed and implemented by faculty, and funded partly by the *National Science Foundation* (NSF); these two disparate approaches to networking will be used repeatedly by other groups. The name is later changed to Computer and Science Network.

Teletel, commonly known as *Minitel* after the terminals that France Telecom gave away throughout France to establish the network, is introduced. Rumor has it the final decision was made by President Mitterand. Minitel was an innovative pioneer in the use of videotext and of national scale information provision. It has grown into probably the largest communicating computer system still not generally exchanging mail with the Internet. No telephone company or national government has since been so bold; although with the independent spread of personal computers, probably none has to be again.

ARPANET has 213 hosts in August 1981.

Early 1980s Various hub and spoke networks such as MAILNET (deceased 1986) spring up to service specialized communities that could not at the time get on ARPANET, access to which was tightly controlled because it was still largely supported by tax dollars.

Reagan administration changes the name of ARPA to DARPA to emphasize its defense role.

1982 *EUnet* (European UNIX Network) begun by Teus Hagen, Peter Collinson, and Keld Simonsen at the April EUUG (European UNIX User's Group) meeting in Paris, to use UUCP to provide mail and USENET news service. EUnet quickly spread throughout Europe.

Tom Jennings invents the FidoNet protocols and FidoNet begins; it originally connects only IBM PCs or compatibles running MS-DOS. Some people involved in the development of FidoNet wanted to use UUCP, but it was thought (erroneously) at the time that the UUCP protocol was proprietary to AT&T, so new protocols were invented. The first Fido transmission protocol was Xmodem with a few simple extensions. The current Fido transfer protocol, ZedZap, is Zmodem with a few simple extensions.

1983 January 1, 1983 is the official date on which all hosts and networks communicating with the ARPANET are required to use TCP/IP instead of the previous NCP protocols. Many hosts don't quite make it by that date, but almost all eventually climb back on.

On the same date, ARPANET splits into two networks. One is *MILNET,* which reverted to more or less operational military uses in the *Defense Data Network* (DDN). The other is still called ARPANET, which becomes the research backbone of what was then called the ARPA Internet.

The Internet was a basically different kind of thing from the ARPANET. The ARPANET was a single network, using a single network protocol. The Internet was an internet, that is a group of many networks using underlying different network protocols and hardware, but all communicating using one common overlying protocol, the Internet Protocol (IP).

4.2BSD, a version of the UNIX operating system, released. This *Berkeley Software Distribution* (BSD) was largely funded by DARPA, which wanted a standard operating system for use by its researchers, including support for TCP/IP networking. Since it was almost entirely supported by funds from the federal government and the state of California, 4.2BSD is made available at cost of distribution.

Inexpensive microprocessors such as the Motorola 68000 and the Intel i286 become available. Old and new computer vendors take advantage of these processors and networking, particularly that provided in 4.2BSD, to produce networked workstations.

EARN, the *European Academic and Research Network,* is established on the model of and interconnected with BITNET.

ARPANET host tables show 562 hosts in August 1983, some elsewhere on the Internet.

1984 Specifications of the *Domain Name System* (DNS) published as RFCs, by Paul Mockapetris, permitting widespread implementation of a distributed host naming scheme.

Bill Gibson publishes a science fiction novel, *Neuromancer,* which describes a worldwide distributed matrix of computer networks supporting a Cyberspace inhabited by humans and programs. Gibson was strongly influenced by previous stories by Vernor Vinge, who was clearly describing successors to the ARPANET. The term, *The Matrix,* seems appropriate to describe all the computer networks worldwide that intercommunicate using at least electronic mail.

continued on next page

History of the Internet *(continued)*

NSF forms an office that can do networking, with the intention of connecting national supercomputer centers.

JANET, the Joint Academic Network, is established in the U.K., partly from the former SERCnet. It uses the *Coloured Book* protocols, which are in some way similar to and influenced both ISO-OSI and TCP/IP.

JUNET, the Japan UNIX Network, established by Jun Murai and others to use UUCP to provide mail and USENET news service throughout Japan.

Mid 1980s The ARPA Internet became the "DARPA Internet" after the name change of its original parent body, then the "Federal Research Internet," to acknowledge the participation and funding of many different U.S. government agencies, including DARPA, DCE, NASA, DoE (Department of Energy), and NSF. Finally, it became known as just the Internet, in acknowledgment of its numerous and diverse participants and sources of funding.

Meanwhile, many widespread distributed networks such as HEPnet, PHYSNET, and MFEnet spread rapidly, using a variety of protocols (HEPnet, DECnet, and MFE, respectively in these examples). Given them plus ISO-OSI, XNS, Coloured Books, and TCP/IP, the future of networking appeared likely to be a plethora of protocols on dissimilar and hardly speaking networks.

1984 NSF expands its interest in networking, and arguments for an *NSFNET* to connect supercomputer centers are published.

1985 U.S. Federal Court Judge Harold Greene breaks up AT&T and permits other companies to provide long distance telephone service. Many companies start laying much fiber optic cable across continents and oceans, providing the basis for later fast data links.

NSF funds the first five national supercomputer centers, as well as a dozen or two regional networks which are intended to interconnect with a national NSFNET backbone. NSF and DARPA agree to permit mutual access of their users to their networks.

PeaceNet, a conferencing system focussed on the peace movement, founded. It later affiliates with like-minded systems through the Association for Progressive Communications, which connects places such as Nicaragua that were at the time politically difficult to reach, and Kenya.

1986 DNS actually does become widespread.

Craig Partridge invents *MX* (Mail Exchanger) records, which permit hosts on non-IP networks such as CSNET and UUCP, and later BITNET and FidoNet, to have domain addresses.

Original NSFNET backbone implemented, using 56Kbps links and "Fuzzball" routers.. It quickly becomes a backbone in the Internet, alongside the ARPANET.

NNTP (Network News Transfer Protocol) specified by Brian Kantor, Phil Lapsley, and others for efficient transmission of USENET news over TCP/IP.

Cleveland Freenet opens service, providing Internet access to the public, and inaugurating a series of related access hosts, many of them banded together in NPTN (National Public Telecomputing Network).

DFN (German Research Network) is established, using ISO-OSI protocols.

Lottor estimates 5,089 Internet hosts in November 1986.

At this point the Internet was large but not completely out of proportion to other networks. BITNET was about the same size in number of users, and UUCP actually had more hosts. The Internet was one among many in the Matrix. Its most popular applications were still the Big Three of TELNET, FTP, and mail, plus some relative newcomers such as USENET news and the X Window System. But it did have TELNET and FTP and other interactive applications, which most other global networks did not.

August 1986 First interoperability workshop organized by Advanced Computing Environments; gathering together much of the TCP/IP industry; invitation-only; 250 attendees. This was the predecessor of the Interop (and later NetWorld+Interop) conferences.

1987 DASnet, a gateway machine that connected diverse networks and conferencing systems, begins operations. This system was considered heretical by many, because it connected both commercial and noncommercial systems, which was at the time just not done.

About this year, WWIVnet, another PC network, modeled partly on FidoNet, begins. This is third level homage to the ARPANET, since WWIVnet was influenced by FidoNet, which was influenced by UUCP and USENET, which were modeled partly after the ARPANET.

NSF agrees for Merit, Inc. to manage the NSFNET backbone, in cooperation with MCI and IBM.

March 1987 First TCP/IP Interoperability Conference; first that doesn't require an invitation; 700 attendees. A 12-page, virtually content-free, "Premiere Issue" of *ConneXions* is published and distributed to all attendees.

May 1987 *UUNET,* the first organization to sell UUCP and USENET access, begins with initial funding from the USENIX Association. Its founder is Rick Adams, who had long been active in USENET, UUCP, and the Internet while running the node `seismo`. UUNET has Internet connectivity from the beginning, and gateways mail and news between UUCP, USENET, and the Internet. UUNET later becomes independent, and also starts providing IP connectivity, using the network name AlterNet.

First real issue of *ConneXions* (Volume 1, No. 1) is published. It contains the now-historic article about ISODE written by Marshall Rose.

December 1987 Second TCP/IP Interoperability Conference; in D.C.; 900 attendees.

1988 NSFNET T-1 (1.544Mbps) backbone operational.

The Morris worm infests the Internet, bringing an anticipatory glare of national and international media attention. Its author is later convicted, breaking with the previous Internet tradition of absorbing crackers, who had to have been good enough to hire. Too many people now depend on the Internet for the old methods to work.

October 1988 Lottor estimates 56,000 hosts on the Internet in October 1988.

U.S. National Bureau of Standards (NBS) publishes *GOSIP* (Government OSI Profile), which is widely misinterpreted as requiring all organizations funded by the U.S. government to switch from TCP/IP to OSI.

continued on next page

History of the Internet *(continued)*

Actually, GOSIP only requires government agencies to buy OSI, not run it. The Internet is still sufficiently dominated by the U.S. government that many people take seriously the concept of converting the Internet to OSI; certain government agencies and other parties try very hard to promote that eventuality.

First INTEROP (new name); Santa Clara; 5,400 attendees.

About 1988 Many factors, including inexpensive microprocessors, the implementation of TCP/IP in the 4.2BSD UNIX operating system, the spread of fiber optic cable, the development of the NSFNET regionals, and the implementation of the T-1 (1.544Mbps) NSFNET backbone, combine to produce an exponential growth rate for the Internet. From being one among many, it quickly becomes the largest of all.

The focus shifts, from government funding to privatization, from resource sharing to resource discovery, from academic and research networks to the beginnings of a global multipurpose network.

1989 Lottor reports 80,000 hosts on the Internet in January 1989.

Berlin Wall falls as Gorbachev and Bush watch. The Cold War ends, but one technology it spawned continues to spread, now extending far beyond any government.

RIPE (Réseaux IP Européens) established as a coordinating body for IP networks in Europe.

PSI (Performance Systems International) incorporated.

NSF specifies its *Acceptable Use Policy* (AUP), which essentially says the Internet Backbone shall only be used for research or education, or in support of research or education. This AUP is widely misinterpreted to mean that the Internet is non-commercial, but it actually applies only to one backbone network in the Internet, and other backbones quickly start appearing.

August 1989 CSNET merges with BITNET under the name *CREN* (Corporation for Research and Education Networking). BITNET continues, but CSNET fades away, as NSFNET and other developments make Internet connectivity readily available.

October 1989 INTEROP in San Jose: 10,000 attendees. Eleven days after this event the Bay Area is struck by a magnitude 7.1 earthquake.

1990 NSFNET T-3 (45Mbps) backbone implementation begins.

ARPANET decommissioned because its hardware and link speeds have become technologically obsolete.

BITNET begins to shrink, as it is absorbed by the Internet. If you were a computer center director, paying $5,000 a year for a link to one BITNET machine, and you discovered all NJE traffic could be carried over the IP link you were already paying for to connect all your other machines on campus, what would you do?

The *archie* indexer of anonymous FTP archives invented at McGill University in Montreal by Alan Emtage and Peter Deutsch. The Internet is suddenly readily usable by people without personal login accounts on more than one system, since archie can find a file for you and anonymous FTP can retrieve it.

WAIS (Wide Area Information Servers) invented by Brewster Kahle, backed of Thinking Machines, Inc., and others. WAIS permits distributed document servers with documents indexed by multiple methods and searched by various techniques.

About this time many library catalogs become accessible through TELNET connections, making anonymous TELNET a useful concept, and further spreading the idea of the Internet as a distributed service provider.

January 1990

Interpolating from Lottor's estimates yields 188,000 Internet hosts in January 1990.

AlterNet begins operations. PSInet begins operations. These are privately owned IP backbone networks of at least national scale.

The World, `world.std.com`, becomes first site to offer full commercial Internet dial-up access to the general public but is restricted to roughly half the net by NSF's AUP.

September 1990

ANS (Advanced Network and Services, Inc.) founded by Merit, MCI, and IBM. ANS effectively runs NSFNET, and also runs ANSNet, which is available for pay through ANS CO+RE.

October 1990

INTEROP in San Jose: 22,000 attendees.

1991: The Year of Europe

Lottor estimates 376,000 Internet hosts in January 1991.

Hosts on the Internet in Europe increase by a factor of four, as Europeans grow tired of waiting for OSI protocols to be specified and implemented and turn to TCP/IP instead.

EUnet decides to commercialize, and incorporates the next year. In addition to its traditional role as a UUCP provider, it becomes the largest IP provider in Europe, connecting dozen member countries, including North African countries such as Algeria and Tunisia and Near Eastern nations such as Turkey and Israel.

In the U.S., Senator Al Gore proposes, the Bush Administration backs, and Congress approves the High Performance Computing Act (HPCA) of 1991, which provides initial funding for a *National Research and Education Network* (NREN). This would be a very high speed (gigabit per second) backbone network in the Internet, with related networking projects.

Gopher invented at the University of Minnesota by Paul Lindner and Mark P. McCahill.

Mike Schwartz of the University of Colorado at Boulder invents *netfind,* for locating people, and the term *Resource Discovery* to refer to it and related protocols such as *archie, WAIS, Gopher,* and *WWW.*

March 1991

Three independent commercial IP connectivity providers, AlterNet, PSINet, and CERFnet, form the *Commercial Internet Exchange* (CIX) and begin transferring traffic among their customers without use of any government funded intermediary. The NSFNET AUP does not apply on traffic exchanged directly among CIX members.

April 1991

U.S. National Institute of Standards (NIST, formerly NBS) publishes GOSIP Version 2. By this time the Internet is so widespread and the government has become a sufficiently small part of it that few people take OSI conversion seriously.

continued on next page

History of the Internet *(continued)*

September 1991 EBONE initiative begins, independently of any government funding, with the intent of building an international IP backbone for Europe.

October 1991 INTEROP in San Jose: 30,000 attendees.

1992: The Year of Gopher Lottor estimates 727,000 Internet hosts in January 1992.

Gopher traffic increases approximately sixfold, leaving traditional services such as FTP behind (in growth rate; not in absolute traffic).

At least two U.S. presidential candidates, including the winning one, use Internet technology extensively in their campaigns.

Poland, Czechoslovakia, and Hungary, followed by Estonia, connect to the Internet, with permission from the U.S. Dept. of Commerce for their traffic to traverse the NSFNET backbone. Later in the year, Russia connects to the Internet through a dialup IP link to a CIX member (AlterNet), and on 24 September 1992 NSF asks and obtains permission for IP traffic from the former Soviet Union to traverse the NSFNET backbone, thus removing the last major Cold War barrier to Internet participation.

OSI is clearly dead, although vestiges of it, such as X.400 (mail) and X.500 (directory) services live on in pockets.

World-Wide Web (WWW) invented by Tim Berners-Lee and others at CERN (European Nuclear Research Center). This protocol and related software permits distributed hypertext across the Internet.

NSFNET T-3 (45Mbps) backbone implementation completed.

January 1992 The Internet Society (ISOC) founded.

August 1992 The World (`world.std.com`) obtains first official permission from NSF to route its Internet dial-up access to NSFnet and thus the entire Internet. This was apparently the first of a new class of retail Internet access providers, or login hosts, that permit users to log in and use Internet services from the provider's machine.

September 1992 EBONE is completed in forming a pan-European IP backbone; this was done as a cooperative venture, without significant initial government backing or funding.

October 1992 INTEROP in San Francisco: 40,000 attendees.

1993: The Year of Mosaic and WWW Lottor estimates 1,313,000 Internet hosts in January 1993.

Bytes transferred with the World-Wide Web (WWW) protocol (as seen crossing the NSFNET backbone) increase by a factor of 2200 (220,000 percent) and packets by 1600 times (160,000 percent) in one year. For comparison, FTP traffic increased approximately 1.8 times, or at about the same rate as the underlying growth rate of hosts on the Internet. Gopher traffic increased approximately 7.4 times in packets and 8.7 times in bytes in 1993, but WWW outpaced it. This increase is presumably mostly due to the *Mosaic* user interface (released in early 1993) to WWW, Gopher, FTP, and other protocols. By providing a common but native interface on DOS, Windows, Macintosh, and UNIX systems to a variety of Internet information resources, Mosaic took WWW from being an also-ran service to the most dynamic on the Internet.

Senator Al Gore becomes Vice President Al Gore. The White House and the House of Representatives connect to the Internet.

DARPA becomes ARPA again as the Clinton administration assigns it a primary role in civilian technology transfer; simultaneously NREN is downsized.

A consortium of industry figures proposes a *National Information Infrastructure* (NII), emphasizing voice and video.

ANS joins CIX, which by now has some two dozen members in many countries worldwide.

PSINet and AlterNet deploy T-3 backbones.

Delphi becomes the first of the large centralized commercial conferencing systems to join the Internet, giving its users access to a wide range of Internet services, not just mail. Some small Gopher servers are saturated because of 100,000 Delphi users using the same menu.

IIKK and AT&T JENS (Spin) begin commercial IP access in Japan, with international IP connectivity to the global Internet. IIJ begins commercial IP access in Japan, and gains international connectivity the following year.

TWICS, a conferencing system begun in 1984, connects to the Internet and starts providing login host service in Tokyo.

August 1993 INTEROP in San Francisco: 55,000 attendees.

October 1993 First TIC/MIDS Internet Demographic Survey estimates a minimum of 741,000 hosts on the Internet. The survey also shows a minimum of 2.5 million Internet users, and a pool of about 5.7 million people with Internet access (whether they use it or not).

1994 Lottor reports 2,217,000 Internet hosts in January 1994.

America Online (AOL), the fastest growing of the centralized commercial conferencing systems, connects to the Internet. The USENET newsgroup `alt.best.of.internet` becomes a laboratory of flaming, due to being first in AOL's list of newsgroups.

Numerous commercial outfits start using the Internet for marketing and sales, as in `florist.com`, the NSTN (Nova Scotia Technology Network) experiment with putting a bookstore on the net, and the Online BookStore.

PSI buys IIKK.

April 1994 WWW passes Gopher in bytes transferred across the NSFNET backbone. This is partly because a greater percentage of the WWW traffic is images, but bytes are bytes.

May 1994 NetWorld+Interop conference draws 65,000 people to Las Vegas; four other NetWorld+Interops are scheduled for the same year, in Berlin, Tokyo, Atlanta and Paris.

June 1994 NetWorld+Interop Berlin; 20,000 attendees.

July 1994 Sprint begins selling IP connectivity.

NetWorld+Interop Tokyo; 40,000 attendees.

September 1994 NetWorld+Interop Atlanta; 60,000 attendees.

CIX board meeting during NetWorld+Interop Atlanta exposes severe internal squabbling. Sprint continues to permit its customers to resell connectivity and IP routing. Net99 announced as an alternative to CIX. Few users notice, since the Internet continues to operate anyway.

continued on next page

History of the Internet *(continued)*

October 1994 EARN and RARE merge to become *TERENA,* the Trans-European Research and Education Networking Association.

Lottor estimates 3,864,000 hosts on the Internet in October 1994.

NetWorld+Interop Paris; 35,000 attendees.

The Second TIC/MIDS Internet Demographic Survey estimates for October 1994 13.5 million users who can use interactive Internet services from 3.5 million hosts in the Consumer Internet in more than 70 countries, from Antarctica to Greenland, from Fiji to Ghana, from Moscow to Cape Town, from Iceland to India. This survey also estimates 7.8 million users of 2.5 million hosts that can provide interactive Internet services in the Core Internet. And it estimates 27.5 million users of electronic mail in the Matrix of all computers that exchange electronic mail.

November 1994 NIST renames GOSIP to POSIT, and recommends TCP/IP.

Republicans win majorities in both House and Senate in U.S. for first time in 40 years, and immediately de-emphasize NSF networking programs even more than Democratic administration had already done. However, so little of the Internet is now supported by the U.S. government that few users notice.

AOL buys ANS.

December 1994 WWW traffic doesn't grow as fast as in 1993, but in 1994 its increase of "only" 300 times on NSFNET is faster than any other protocol, and by the end of the year WWW is second only to FTP in traffic on NSFNET. Exact amount of WWW traffic on the Internet at large is unknown, but clearly has grown at a similar rate.

CIX continues to self-destruct through internal squabbling, as it has done all year. Few users notice, since the Internet continues to operate anyway.

BBN buys SURAnet, adding it to BARRnet and NEARnet in its collection of Internet regional networks.

January 1995 Lottor estimates 4,852,000 hosts on the Internet in January 1995.

Brian Reid estimates 16.5 million USENET users in January 1995. USENET isn't the Internet, but it covers much of it, and Reid's figure is slightly larger than the TIC/MIDS estimate of Internet users for October 1994, just as would be expected.

Microsoft buys shares in UUNET, the parent of AlterNet. MCI begins offering Internet access.

Prodigy permits its users interactive WWW access and gets several hundred thousand takers almost immediately.

CompuServe permits some interactive Internet access to its users, and announces PPP service to begin in May.

All three of the biggest centralized conferencing systems, AOL, CompuServe, and Prodigy, now permit some interactive access to the Internet, but none of them provide the full array of services almost any true *Internet Service Provider* (ISP) supports for its customers.

Meanwhile, true ISPs number in the hundreds, if not thousands, and are collectively worth many millions of dollars.

Summary

Some themes seem clear in hindsight.

Government funding of network protocol development and network deployment is very useful. Participation by ARPA, DCA, NSF, NASA, and other government agencies in development of TCP/IP and the Internet is clearly a major reason that the United States currently has 60% of all users, hosts, and networks on the Internet, and the TCP/IP industry of routers and other hardware and software is primarily based in the United States.

But governments have limited foresight. The original ARPANET plan did not include communication services such as mail, lists, or news. People who wanted these things invented them. The same thing happened with *archie, Gopher,* and *Mosaic.*

Many of the networks mentioned in this timeline have come and gone, but they all played their roles in producing the networking environment we have today. Others have survived and thrived. USENET has become a worldwide distributed conferencing system. UUCP and FidoNet still go first and most inexpensively into many places. And the Internet has become so large and obviously successful, connecting as it does millions of computers and tens of millions of people in more than 80 countries worldwide, that it can't be ignored. Even the big centralized conferencing services are all joining the Internet, while hundreds of smaller, more sophisticated Internet Service Providers grow like weeds.

Sources

[1] Early ARPANET host counts and date for ARPANET electronic mail from Peter Salus `<phs@netcom.com>`; see his book *Casting the Net,* Addison-Wesley, 1995.

[2] Later Internet host estimates from Mark Lottor `<mkl@nw.com>`

[3] Internet growth rates and demographic survey information from Matrix Information and Directory Services, Inc. `<mids@mids.org>`

[4] NSFNET byte and packet counts from Merit, Inc. Available from: `ftp://nic.merit.edu/nsfnet/statistics`

[5] USENET user estimate from Brian Reid, as posted in USENET newsgroup `news.lists`

[6] ITS information from Michael A. Patton `<MAP@BBN.COM>`

[7] FidoNet details from Eric McKinney `<ericm@tic.com>`

[8] The World history from Barry Shein `<barry@world.std.com>`

[9] Gopher history from Peter Deutsch `<peterd@bunyip.com>`

[10] Most other material is from the book, *The Matrix: Computer Networks and Conferencing Systems Worldwide,* Digital Press, 1990, and from articles in *Matrix News.*

We are interested in any suggestions you may have for updates; please mail them to: `tic@tic.com`.

JOHN S. QUARTERMAN is a partner in Texas Internet Consulting (TIC) `<tic@tic.com>`. He is a partner in Zilker Internet Park `<info@zilker.net>`, which provides Internet access from Austin, Texas. He is Editor of the monthly newsletter *Matrix News* and the color *Matrix Maps Quarterly,* both published by Matrix Information and Directory Services, Inc., (MIDS) `<mids@mids.org>`. He has written or co-authored six books: *The Matrix,* Digital Press, 1990; and *The Design and Implementation of the 4.3BSD UNIX Operating System,* 1989, *UNIX, POSIX, and Open Systems,* 1993, *Practical Internetworking with TCP/IP and UNIX,* 1993, *The Internet Connection,* 1994, and *The E-Mail Companion,* 1994, all from Addison-Wesley. He can be reached as: `jsq@tic.com`

Security Activities at the Winter IETF Meeting

by Jim Galvin, Trusted Information Systems

Introduction

The San Jose *Internet Engineering Task Force* (IETF) meeting, held December 5–9, 1994 set a new record for attendance with over 1,000 attendees. Security discussions were prevalent everywhere, including the terminal room, which was subject to an active attacker who, with limited success, tried to keep folks from using the network. (The motives of the attacker are unknown, but the occasional unavailability of the network was immediately apparent to everyone in the room.) Following are summaries of the significant security-relevant activities of the week.

Updates from previous meeting

Telnet Encryption Option: At the previous IETF, held in Toronto during the week of July 25–29, 1994 it had been agreed that a document should be produced specifying the existing Telnet encryption option practice. [7] However, a serious vulnerability has since surfaced in the existing practice; a better version of the protocol is needed. The Security and Applications Area Directors will work to create a design team and/or working group to focus on getting a better version of the Telnet Encryption Option documented.

Key Management Working Group: Also at the previous meeting, it had been agreed that a key management working group would be created. However, since that time the *Internet Protocol Security* (IPSEC) Working Group has developed a proposal for a session key establishment protocol. Therefore, this action item has been subsumed by the IPSEC Working Group.

DNS Security

Consensus was reached that the *Domain Name System* (DNS) Security Working Group should proceed with the Eastlake/Kaufman proposal (one of two that had been under consideration) for adding digital signature services to the DNS protocol. Some issues, pertaining to topics such as revocation, choice of algorithm, and the signing of nonauthoritative data, were discussed and a revised specification is expected shortly. Depending on implementation experience, the proposal may be submitted for publication as a proposed standard prior to the next IETF.

IP Security

Significant progress on IP security was made during this meeting. There appears to be rough consensus on the IP security protocol proposal. In addition, in spite of the several key management alternatives presented, all participants appear willing to work together to arrive at a consensus proposal that meets the requirements of the working group. Finally, it is believed that all known patent issues have been resolved.

PEM

The consensus of the *Privacy Enhanced Mail* (PEM) Working Group was to submit the two MIME-PEM integration documents for publication as proposed standards. A last call will be issued on the mailing list to allow folks not in attendance to comment. The first document (Security Multiparts for MIME: Multipart/Signed and Multipart/Encrypted) specifies a framework for support of security services in MIME. The second document (PEM Security Services and MIME) is a specification of how the PEM services use the framework to support the application of digital signatures and encryption to MIME body parts.

New call for Action

The Router Requirements Working Group, which was inactive for the last few years due to lack of a volunteer to edit the document, has been resurrected.

Fred Baker (of Cisco Systems) has volunteered to be the editor of the new version of the document, which has been posted and is currently under discussion and review. The earlier version of the Router Requirements document, which had been languishing in the working group for years, has been published and designated historical.

New working groups

Two new working groups will be formed in the security area: *Hyper-Text Transfer Protocol* (HTTP) Security and Object/Document Security. The first will address providing security services for the HTTP/*World-Wide Web* (WWW) suite while the second will focus on providing services for documents and other "static" objects. Both of these groups met as birds of a feather sessions at this IETF. Both are expected to draft charters for approval before the next IETF meeting.

More information

The next IETF meeting is scheduled for April 3–7, 1995, in Danvers, Massachusetts. For more information about the IETF security area, join the `saag@tis.com` e-mail list by sending a subscription request message to the `saag-request@tis.com` address.

For more information about the IETF, check out its Web home page: `http://www.ietf.cnri.reston.va.us`, which includes pointers to all working groups, RFCs, Internet Drafts, charters, mailing lists, and mailing list archives.

References

[1] Dern, D., "Interview with Steve Kent on Internet Security," *ConneXions,* Volume 4, No. 2, February 1990.

[2] Galvin, J., "The Deployment of Privacy Enhanced Mail," *ConneXions,* Volume 5, No. 10, October 1991.

[3] Galvin, J., "Components of OSI: The Security Architecture," *ConneXions,* Volume 4, No. 8, August 1990.

[4] Schiller, J., "Issues in Internet Security," *ConneXions,* Volume 7, No. 9, September 1993.

[5] *ConneXions,* Volume 4, No. 8, August 1990, "Special Issue on Network Management and Network Security."

[6] Galvin, J., "Security Awareness Increasing within the IETF," *ConneXions,* Volume 8, No. 9, September 1994.

[7] Galvin, J., "Summer IETF Meeting Security Report," *ConneXions,* Volume 8, No. 10, October 1994.

[8] Stallings, W., "Cryptographic Algorithms, Part I: Conventional Cryptography," *ConneXions,* Volume 8, No. 9, September 1994.

[9] Stallings, W., "Cryptographic Algorithms, Part II: Public-Key Encryption and Secure Hash Functions." *ConneXions,* Volume 8, No. 10, October 1994.

[10] Stallings, W., "Pretty Good Privacy," *ConneXions,* Volume 8, No. 12, December 1994.

[Ed.: A version of this report appeared in the *Data Security Letter,* January 1995. For more info contact: `dsl@tis.com`]

JAMES M. GALVIN is a Senior Computer Scientist at Trusted Information Systems (TIS). Dr. Galvin's responsibilities emphasize communications security, especially computer networks, architectures, policies, and procedures. He is a principal in the development of TIS' openly available implementation of Privacy Enhanced Mail. He is very active in the IETF Security Area, serving as Executive Director of the Security Directorate, and past Chair of the OSI Implementor's Workshop Security Special Interest Group, hosted quarterly by the National Institute of Standards and Technology. He received his Ph.D. and M.S. degrees, both in Computer Science, from the University of Delaware in 1988 and 1986, respectively. In 1982, he received his B.S. in Computer Science and Mathematics from Moravian College in Bethlehem, PA. E-mail: `galvin@tis.com`

Book Reviews

ATM Networks: Concepts, Protocols, Applications (second edition) by Rainer Händel, Manfred N. Huber and Stefan Schröder, Addison Wesley, 1994. ISBN 0-201-42274-3.

General treatment

There are probably as many books on Asynchronous Transfer Mode (ATM) and Broadband ISDN (Integrated Services Digital Network) as there are bytes in the ATM cell user payload. Some are highly specialized: Raif Onvural's *Asynchronous Transfer Mode Networks: Performance Issues* is representative of this genre. Händel, Huber, and Schröder's production falls into what I would consider the "general treatment of ATM" category. The first 3 chapters provide a crisp and well articulated overview of ATM and Broadband ISDN. Chapter 4 and early sections of Chapter 5 do bog down a bit in standards-ese and acronymania (it's ATM-azing how many acronyms one can squeeze into a 250-page book on a data link technology...), but I'd encourage you to read on. The authors provide a highly accurate and succinct description of ATM that will be useful for those who want to understand both general and user-network aspects of this technology.

Good explanations

The strength of the book lies in the latter sections of Chapter 5, where the authors describe the "higher layers and interworking of the [ATM] user plane." Here the authors do a fine job of explaining the roles of ATM as a bearer service for Frame Relay and SMDS, providing a much-needed accurate perspective of the relationship between these services and ATM. A concise description of ATM signalling is provided in chapter 6 (a minor criticism here is that the authors assume the reader is familiar with Narrowband ISDN signalling). Chapter 7 is an excellent overview of ATM switching: again, succinct and highly readable, with good comparisons of different switching and networking techniques. I commend the authors for having done such a fine job of defining and comparing matrix switching against central memory switching, and multi-path networking against single-path (Banyan) networks. Chapters 8 and 9 describe ATM transmission networks and the evolution paths to a broadband ISDN environment, and chapters 10 and 11 cover miscellany: issues for voice over ATM, tariffing, telecommunications management networks, etc. While these are generally good chapters, the authors might have included descriptions of more recent trials or operational ATM networks, but this is again a minor criticism.

An easy read

In certain respects, there is little that distinguishes the material covered in this book from that covered in Martin DePrycker's popular and successful *ATM Transfer Mode: Solution for Broadband ISDN* (second edition, Ellis Horwood Ltd., 1993, ISBN 0-13-178542-7). Where the books indeed differ is in the level of presentation: whereas DePrycker's book is a highly analytical treatment of ATM, replete with nearly every mathematical formulae necessary to describe queueing and switching strategies associated with ATM, Händel, et. al., do a commendable job of presenting ATM in clean and concise fashion. For a technical book, it is "an easy read."

Having spent a good deal of time on both the computer communications and carrier network ends of the telecommunications spectrum, I would definitely place this book on the telephony networking shelf of my bookcase. The discussions of telephony issues—performance, operations, administration and management, quality of service, telephony transmission systems characteristics—are more detailed and elaborate than discussions about LANs and internetworking: the authors might consider local ATM, multiprotocol and routing operation over ATM in a future edition or revision.

Gratefully, however, the authors do not use X.25 as the lone example of data networking. For the presumed audience of *ConneXions,* the middle chapters will prove to be the most informative and useful.

Well-written introduction

If you are looking for a solid, sweeping introduction to ATM, this book is a good choice. It's also a good choice if you simply enjoy reading books that are extremely well-written: I find it simultaneously remarkable and disappointing that the authors, decidedly European, have a better grasp and respect for the English language than those for whom English is the native language.

—David M. Piscitello,Core Competence, Inc.

The E-Mail Frontier—Emerging Markets and Evolving Technologies by Daniel J. Blum and David M. Litwack, Addison-Wesley, 1994. ISBN 0-201-56860-8.

Most books on a technology simply cover the technical details. Few try to cover an entire "industry" roaming the full terrain of technology and business. This is one of those efforts, looking at an industry which shows the peculiar characteristics of being twenty years old, yet oddly nascent in much of its service and product offerings. Unless the book were entirely incompetent, I would automatically recommend it by virtue of its scope and the importance of the industry. In fact, the book is quite competent.

X.400 and Internet mail

Truth in packaging: I was a compensated pre-publication reviewer for this book, with respect to its discussion of Internet e-mail and the Internet standards process. I'd like to claim that this makes the book's coverage of those topics stellar but it isn't quite that good. The authors have extensive background in electronic mail consulting. Such a business is dominated by work with proprietary products and public messaging services, with interconnection dominated by X.400 technology. Until recently, there was essentially no "Internet e-mail" industry. Therefore like most workers in the commercial sector of e-mail, the authors had limited contact with the Internet. The book reflects that limitation, though not egregiously. In fact, the authors do an excellent job of attempting to overcome what one might call bad training. They do this by focussing on pragmatics.

The usual failing of e-mail discussions is the mystical hand-wave that promises the ultimate success of X.400, ignoring its astonishingly slow growth and the real and massive traffic dominance by Internet e-mail. After more than 10 years, one must wonder at the style of thinking that allows one to ignore the power of a huge installed base, such as the Internet.

Comprehensive introduction

The E-Mail Frontier is broad, thorough and well-organized. It discusses business, marketing, product, service, technology, and standards issues, attempting to provide a pragmatic view of the entire industry. It largely succeeds. Don't rely on it as a definitive source for particular details. Instead, use it as a comprehensive introduction and a broad analysis of the messaging industry. If you are a communications manager, are involved in product marketing, or are otherwise concerned with global, distributed applications services, you should read *The E-Mail Frontier.*

—Dave Crocker, Brandenburg Consulting

Call for Papers

The *ACM-Springer Multimedia Systems* journal is preparing a Special Issue on *Multimedia and Multisensory Virtual Worlds* for publication in December 1995

Background

Virtual worlds are going beyond 3D graphics and are beginning to use multimedia and multisensory technologies such as video, spatial sound, speech, images, haptic and tactile feedback, and wind and heat sensation. This has led to new applications for virtual worlds in science, engineering, medicine, business, training, entertainment and arts to explore physical environments that exist remotely (telepresence), or simulated environments that do not or could not exist; to enrich existing environments (augmented realities); and to develop physical analogues for abstract quantitative and organizational data.

Topics

Original, unpublished research and practice and experience papers are sought that address issues in the design, implementation, and evaluation of virtual worlds that use multimedia and multisensory technologies. Topics include, but are not limited to:

- Multimedia and multisensory interfaces for virtual worlds
- Software architectures for using multimedia in virtual worlds
- Enhancing presence with multimedia and multisensory technologies
- Distributed and multi-user virtual worlds
- Knowledge-based multimedia world modeling
- Manual and automated multimedia world design facilities
- Navigation, search, and retrieval in large multimedia virtual worlds
- Novel applications in visualizing, exploring and manipulating rich multimedia information spaces
- Evaluation of the effectiveness of multimedia virtual worlds, and their impact on users, applications, and organizations

Submissions

Five (5) copies of each manuscript should be submitted to the special issue editor at the address below. For papers that do not include color pictures, e-mail submission is encouraged.

Gurminder Singh
Institute of Systems Science
National University of Singapore
Kent Ridge
Heng Mui Keng Terrace
Singapore 0511
REPUBLIC OF SINGAPORE
Phone: +65 772-3651
Fax: +65 774-4998
E-mail: `gsingh@iss.nus.sg`

Important dates

Submissions due:	May 15, 1995
Notification of acceptance:	August 31, 1995
Revisions due:	September 31, 1995
Publication:	December, 1995

Letter to the Editor

Ole:
Just had first occasion to use the brand-new 1987–1994 index to locate background material on IPng for some big presentations I'm giving in Japan soon. What a time saver. Because I have an efficent person sorting/storing all old issues (of everything) it took about three minutes, solo, late on Sunday night. PERFECT.

Cheers,
—*Bruce Nelson, Auspex Systems*

Bruce:
Glad you like it. Remember that you can also get a cumulative index in ASCII from our `www.interop.com` server. This will let you perform searches using your favorite grepping tool. If you don't have Web access (or find the response too slow—let's face it this does happen from time to time), simply send e-mail to `connexions@interop.com` asking for the ASCII version of the index, and we will send it to you right away.

—*Ole*

Write to *ConneXions*!

We'd love to hear your comments, suggestions and questions about anything you read in *ConneXions*. Our editorial address is given below. Use it for letters to the Editor, questions about back issues etc.:

ConneXions—The Interoperability Report
303 Vintage Park Drive, Suite 201
Foster City, CA 94404–1138
USA
Phone: +1 415-578-6900 or 1-800-INTEROP (Toll-free in the USA)
Fax: +1 415-525-0194
E-mail: `connexions@interop.com`
URL: `http://www.interop.com`

Subscription information

For questions about your subscription please call our customer service hotline: 1-800-575-5717 or +1 502-493-3217 outside the USA. This is the number for our subscription agency, the Cobb Group. Their fax number is +1 502-491-8050. The mailing address for subscription payments is P.O. Box 35840, Louisville, KY 40232–9496.

NetWorld+Interop 1995 World Tour

1995 NetWorld+Interop events

Here are the dates and locations for our 1995 NetWorld+Interop events. Check out our home page at `http://www.interop.com` for details:

NETWORLD+INTEROP'95

Frankfurt, Germany	*May 29–June 2*
Tokyo, Japan	*July 17–21*
Paris, France	*September 11–15*
Atlanta, Georgia	*September 25–29*

CONNEXIONS
303 Vintage Park Drive
Suite 201
Foster City, CA 94404-1138
Phone: 415-578-6900
FAX: 415-525-0194

FIRST CLASS MAIL
U.S. POSTAGE
PAID
SAN JOSE, CA
PERMIT NO. 1

ADDRESS CORRECTION
REQUESTED